Data Visualization: CO2 Concentration

Harris Policy Labs, Spring 2020

Susan Paykin

4/26/2020

Introduction

This document shows data preparation and visualization of yearly carbon dioxide concentration data for 24 African countries from 2015-2018 (all available dates). These countries were selected based on research project scope that focuses on data available via Afrobarometer dataset. All data was downloaded courtesy of AidData, a Research Lab at William & Mary.

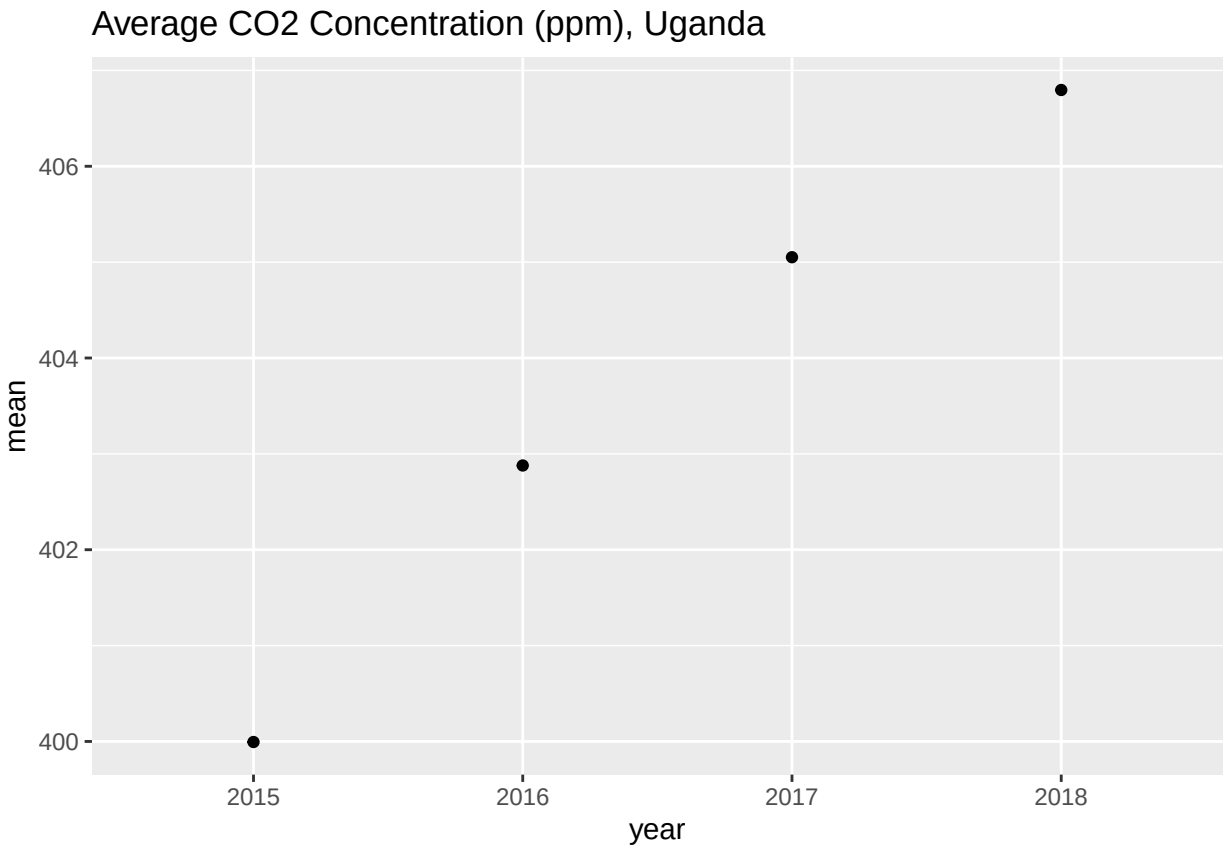
Original data source: O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99-121, https://doi.org/10.5194/amt-5-99-2012, 2012.

Data Preparation

Test: Uganda

As an initial preparation step, I tested the data for a single country, Uganda, to ensure I would be able to manipulate and visualize the approriate data points, as well as test and troubleshoot my initial code.

```
setwd("~/Desktop/Spring 2020/Policy Lab")

#Load data
uganda_co2 <- read.csv("Data/Data Viz Assignment/5ea35212c15e005bdf603974/5ea35212c15e005bdf603974_resul

# Test - Aggregate Uganda data
UGA_national <-
  uganda_co2 %>%
  filter(iso == "UGA") %>%
  summarize(country = "UGA",
            "2015" = mean(oco2_xco2_yearly.2015.mean),
            "2016" = mean(oco2_xco2_yearly.2016.mean),
            "2017" = mean(oco2_xco2_yearly.2017.mean),
            "2018" = mean(oco2_xco2_yearly.2018.mean)) %>%
  gather("year", "mean", -country)

# Plot initial Uganda data
ggplot(UGA_national, aes(x = year, y = mean)) +
  geom_point() +
  ggtitle("Average CO2 Concentration (ppm), Uganda")
```



All Afrobarometer countries

With the Uganda data running smoothly, I loaded the rest of the subnational country data. Each country's data is available in varying levels of subnational geoboundaries, ranging from ADM 1 to ADM 4. Due to the nature of this exercise in which we want to observe trends across countries for select variables, I decided that it would be most appropriate to aggregate the subregional data to the national level for plotting. For this I created a new dataset, `national`, that is composed of the mean CO2 concentrations by year for each country.

```
setwd("~/Desktop/Spring 2020/Policy Lab")

# Read in all Afrobarometer country data
files <- list.files(path = "Data/Data Viz Assignment/subnational_data/", pattern = "*.csv", full.names =
co2 <- sapply(files, read_csv, simplify = FALSE) %>% bind_rows(.id = NULL)

# Aggregate mean CO2 concentrations by year, country
national <-
  co2 %>%
  group_by(iso) %>%
  summarize("2015" = mean(oco2_xco2_yearly.2015.mean),
            "2016" = mean(oco2_xco2_yearly.2016.mean),
            "2017" = mean(oco2_xco2_yearly.2017.mean),
            "2018" = mean(oco2_xco2_yearly.2018.mean)) %>%
  gather("year", "mean", -iso)
```

Global data

I also wanted to compare countries' average concentration levels with the global average, so I identified and loaded this data separately.

```
# Load global CO2 data
library(jsonlite)
json_file <- 'https://datahub.io/core/co2-ppm-daily/datapackage.json'
json_data <- fromJSON(paste(readLines(json_file), collapse=""))

path_to_file = json_data$resources$path[i]
data <- read.csv(url(path_to_file))

# Aggregate mean CO2 concentrations by year
global <-
  data %>%
  mutate(year = year(date)) %>%
  group_by(year) %>%
  filter(year == c("2015", "2016", "2017", "2018")) %>%
  summarise(mean = mean(value))

# Wite as CSV
write.csv(global, "global.csv")
```

```
# Upload global csv
global <- read.csv("global.csv")
```

Visualizations

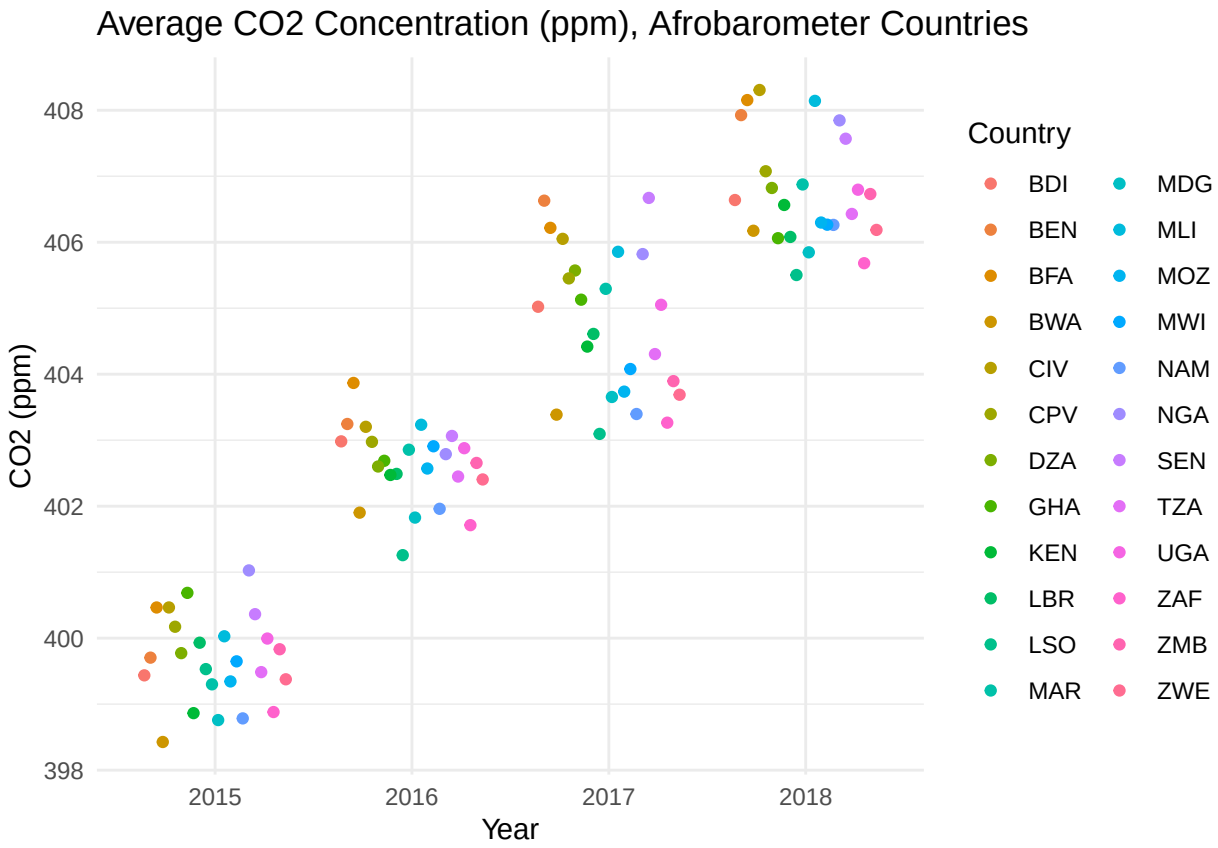
I created three plots to visualize the data in slightly different ways: (A) scatterplot showing the exact points, (B) line graph with conneting trend lines, and (C) line graph with global annual mean trend line for comparison.

This first code chunk sets up the color pallette and the off-set position for the data points to ensure readability.

```
# Set colors for all countries
colourCount = length(unique(national$iso))
getPalette = colorRampPalette(brewer.pal(9, "BrBG"))

# Set position
pd <- position_dodge(0.75)
```

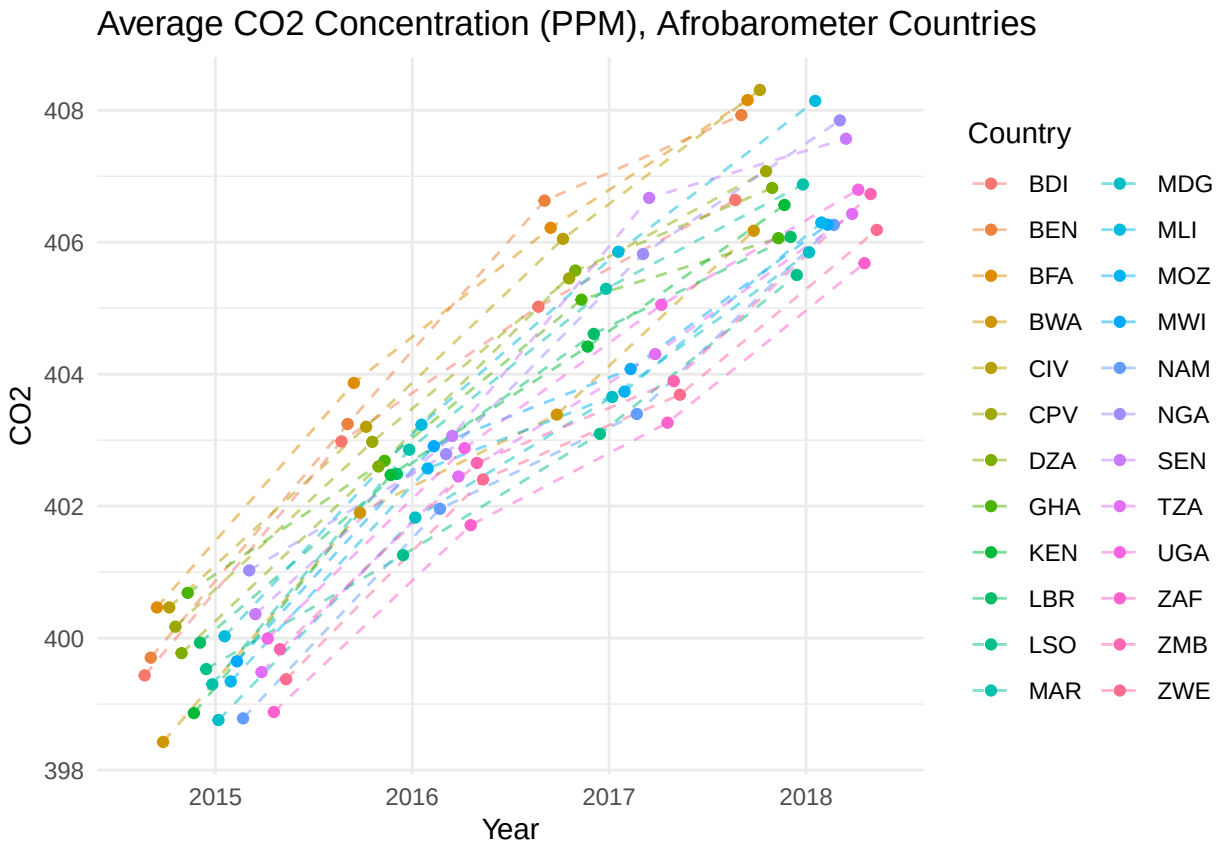
Scatterplot

```
scatter <- ggplot(national, aes(x = year, y = mean, color = iso, group = iso)) +
  geom_point(aes(color = iso), position = pd) +
  scale_fill_manual(values = getPalette(colourCount)) +
  theme(legend.position = "bottom") +
  xlab("Year") +
  ylab("CO2 (ppm)") +
  ggtitle("Average CO2 Concentration (ppm), Afrobarometer Countries") +
  labs(color = "Country") +
  theme_minimal()

scatter
```



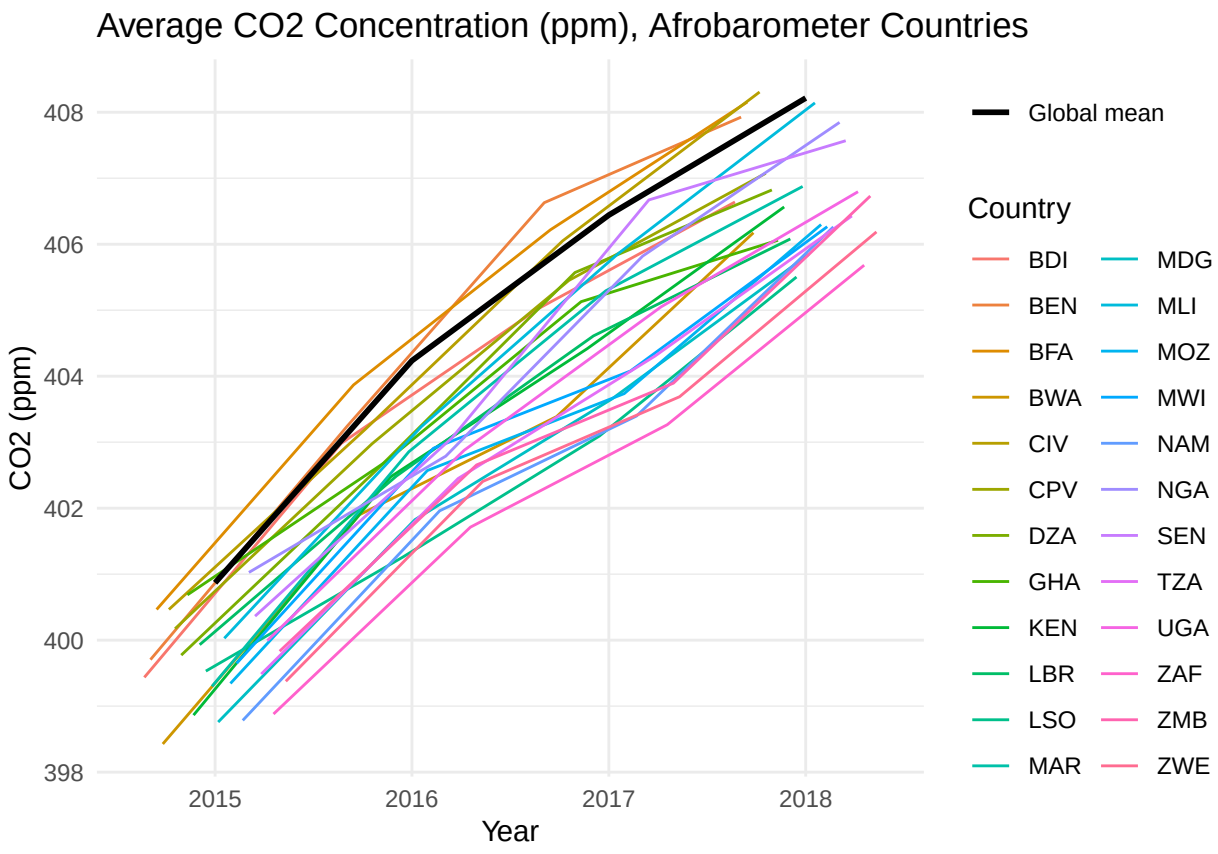
Trend line

```
line_points <-
  ggplot(national, aes(x = year, y = mean, color = iso, group = iso)) +
  geom_line(aes(color = iso), position = pd, linetype = "dashed", alpha = 0.5) +
  geom_point(aes(color = iso), position = pd) +
  scale_fill_manual(values = getPalette(colourCount)) +
  theme(legend.position = "bottom") +
  xlab("Year") +
  ylab("CO2") +
  ggtitle("Average CO2 Concentration (PPM), Afrobarometer Countries") +
  labs(color = "Country") +
  theme_minimal()

line_points
```



Line graph with global mean trend line

```
# Change global year var to character
global$year <- as.character(global$year)

# Plot with global trend line
line_global <-
  ggplot() +
  geom_line(data = national, aes(x = year, y = mean, color = iso, group = iso), size = 0.5, position = p
  geom_line(data = global, aes(x = year, y = mean, linetype = "Global mean"), group = 1, size = 1.0, pos
  scale_fill_manual(values = getPalette(colourCount)) +
  xlab("Year") +
  ylab("CO2 (ppm)") +
  ggtitle("Average CO2 Concentration (ppm), Afrobarometer Countries") +
  theme_minimal() +
  labs(color = "Country") +
  labs(linetype = "")

line_global
```


Average CO2 Concentration (ppm), Afrobarometer Countries
CO2 (ppm)
408
406
404
402
400
398
2015
2016
2017
2018
Year
Global mean
Country
BDI
BEN
BFA
BWA
CIV
CPV
DZA
GHA
KEN
LBR
LSO
MAR
MDG
MLI
MOZ
MWI
NAM
NGA
SEN
TZA
UGA
ZAF
ZMB
ZWE